

M. Roszkowski *Rola dziedzinowych systemów hipertekstowych w zarządzaniu informacją w nauce. Referat wygłoszony podczas konferencji pt. „Zarządzanie informacją w nauce” w Katowicach w dn. 23-24.11.2006 r.*

mgr Marcin Roszkowski
Akademia Świętokrzyska w Kielcach
Filia w Piotrkowie Trybunalskim
Biblioteka

Rola dziedzinowych systemów hipertekstowych w zarządzaniu informacją w nauce.

Dziedzinowe systemy hipertekstowe od końca lat 90-tych ubiegłego wieku są ważnym internetowym narzędziem informacyjnym dla sektora nauki i szkolnictwa. Istotą funkcjonowania tego rodzaju systemów jest gromadzenie i udostępnianie odesłań hipertekstowych do wyselekcjonowanych zasobów Internetu z danej dziedziny wiedzy [Roszkowski 2005; Derfert-Wolf 2004]. O włączeniu do zbioru informacyjnego decyduje zarówno zawartość merytoryczna źródła jak i jego cechy formalne. Dla każdego z systemów opracowywana jest polityka gromadzenia źródeł informacyjnych, zasady odwzorowywania ich cech treściowych i formalnych, a także architektura informacji zakładająca odpowiednią organizację informacji. Celem tego rodzaju systemów jest udostępnianie użytkownikom na wyjściu z systemu, metainformacji w postaci charakterystyk wyszukiwawczych dokumentów elektronicznych opublikowanych w Internecie. Realizacja dziedzinowego modelu udostępniania informacji polega na zawężeniu zakresu zbioru informacyjnego systemu zazwyczaj do jednej dziedziny wiedzy lub działalności ludzkiej. Dziedzinowe systemy hipertekstowe są efektywnym narzędziem informacyjnym, które najczęściej jest wykorzystywane w sektorze szkolnictwa wyższego. Rola tego rodzaju systemów w zarządzaniu i transferze informacji do sektora nauki zostanie przedstawiona na przykładzie działającego w Wielkiej Brytanii systemu Intute

System Intute (www.intute.ac.uk) uruchomiono 13 sierpnia 2006 roku [<http://www.intute.ac.uk/news/71.html>]. Jest on jednym z największych aktualnie funkcjonujących w Internecie dziedzinowych systemów hipertekstowych. Geneza systemu związana jest pośrednio z projektem *eLib*, który ukonstytuował model dziedzinowego dostępu do informacji o zasobach Internetu [Dempsey, Law 2000]. Efektem tego programu było również wytyczenie kierunku badań nad tego rodzaju systemami, co przyczyniło się do wzrostu liczby nowopowstających systemów tego typu, szczególnie na terenie Wielkiej Brytanii. Tym samym, pod koniec lat dziewięćdziesiątych ubiegłego stulecia pojawiła się potrzeba powstania instytucji, która koordynowałaby prace nad wdrażaniem nowych systemów oraz sprawowała opiekę merytoryczną nad już istniejącymi. Funkcję tę miała pełnić sieć o nazwie *Research Discovery Network* (RDN). Celem powstania RDN było oprócz wspomnianych przesłanek, rozwiązanie kwestii źródeł finansowania funkcjonujących systemów, które działały jako projekty wielu instytucji naukowych w Wielkiej Brytanii. Powstała w 1997 sieć RDN [Hiom 2006]

zrzeszała 8 dziedzinowych systemów hipertekstowych o różnych, rozłącznych zakresach przedmiotowych. Były to:

- Artifact – sztuka oraz tzw. kreatywne gałęzie gospodarki (sztuka, architektura, media, muzyka, moda, itp.)
- EEVL – inżynieria, matematyka, informatyka,
- GeSource – geografia
- BIOME – medycyna i nauka o zdrowiu,
- ALTIS – sport, turystyka, spędzanie wolnego czasu,
- Humbul – nauki humanistyczne,
- PSIGate – fizyka,
- SOSIG – szeroko rozumiane nauki społeczne.

Sześć lat funkcjonowania sieci spowodowało, że systemy które zrzeszała stały się wiarygodnymi źródłami informacji dla sektora nauki. Świadczyły o tym zarówno wysokie statystyki odwiedzin systemów, jak również duża liczba wyselekcjonowanych i skatalogowanych zasobów Internetu tworzących ich zbiory informacyjne.

Dynamicznie ewoluująca wirtualna rzeczywistość wymogła na dziedzinowych systemach hipertekstowych konieczność zmian zarówno w architekturze systemów jak i w koncepcji tego rodzaju usług. Tym samym, niezbędnym stało się przeformułowanie zasad funkcjonowania sieci i skierowanie uwagi w jej rozwoju na model zintegrowanej platformy zaspokajającej potrzeby informacyjne wybranej grupy użytkowników i dostosowanej do jej wymagań. Odpowiedzią ze strony środowiska akademickiego Wielkiej Brytanii był system Intute - wysoce kompleksowy system informacyjny oparty na modelu dziedzinowego systemu hipertekstowego. System Intute gromadzi informacje o ponad 14 tysiącach dokumentów elektronicznych i obiektów informacyjnych w Internecie (dane z dn. 05-11-2006 r.). Jest on zbudowany z czterech podsystemów, których zakresy wyznaczają cztery grupy tematyczne:

- nauka i technologia – *Intute Science & Technology* (<http://www.intute.ac.uk/sciences/>)
- sztuka i nauki humanistyczne – *Intute Arts & Humanities* (<http://www.intute.ac.uk/artsandhumanities/>)
- nauki społeczne – *Intute Social Sciences* (<http://www.intute.ac.uk/socialsciences/>)
- medycyna i nauka o zdrowiu – *Intute Medicine & Health Sciences* (<http://www.intute.ac.uk/healthandlifesciences/>)

Pracę systemu koordynuje konsorcjum, w którego skład wchodzi przedstawiciele siedmiu uniwersytetów z Wielkiej Brytanii. Za finansowanie systemu odpowiadają Joint Information Systems Committee (JISC), Arts and Humanities Research Council (AHRC) oraz Economic and Social Research Council (ESRC).

System Intute skierowany jest do użytkowników wywodzących się z sektora szkolnictwa wyższego w Wielkiej Brytanii (HE – ang. higher education) oraz kształcenia na poziomie poniżej szkolnictwa wyższego dla osób, które przekroczyły wiek obowiązku szkolnego (FE – ang. further education). Istota systemu, polegająca m.in. na gromadzeniu informacji o wysokiej jakości zasobach Internetu, pozwala na stwierdzenie, iż stanowi on repozytorium metadanych dla wirtualnych środowisk edukacyjnych (VLE – ang. virtual learning environments). Charakter potrzeb informacyjnych jego potencjalnych użytkowników określa warunki, jakie spełniać muszą elementy jego zbioru informacyjnego. Wiąże się to również z dostosowaniem formatu opisu tego rodzaju zasobów informacyjnych. Dokument elektroniczny, czy też obiekt informacyjny traktowany jest tutaj jako potencjalny obiekt edukacyjny. Dlatego też jego charakterystyka treściowa i formalna musi odwzorowywać jego cechy relewantne w stosunku do specyfiki grupy użytkowników.

Istotną cechą systemu jest wykorzystanie standaryzowanego schematu metadanych do odwzorowania treści i formy obiektów informacyjnych, które stanowią jego zbiór informacyjny. Format opisu został dostosowany do identyfikacji relewantnych cech opisywanych obiektów. Struktura rekordów bazy danych zakłada funkcjonowanie elementów różnych schematów metadanych, których celem jest efektywna charakterystyka źródła internetowego. Format opisu zastosowany w systemie Intute opracowano na podstawie schematów:

- Dublin Core,
- RLLOMAP – profil aplikacyjny RDN/LTSN Learning Object Metadata,
- AACR2 – Anglo American Cataloging Guidelines wyd. 2,
- RDA - Resource Description and Access. [Knowlton 2006, CHAMPAN 2006]

Charakterystyka elementów zbioru informacyjnego jako potencjalnych obiektów edukacyjnych (LOM – ang. learning object metadata) [Najjar, Duval 2006] zakłada wyodrębnianie cech istotnych z dydaktycznego punktu widzenia. W przypadku systemu Intute określono kilka dodatkowych cech dokumentu, którym, podczas indeksowania przypisuje się wartości pobierane z list kontrolowanych. Są to m.in.,

Audience - określenie grupy użytkowników docelowych,

Educational Level - określenie poziomu dydaktycznego dokumentu

Zastosowanie odpowiednich schematów metadanych umożliwia dostęp do informacji za pomocą efektywnych mechanizmów wyszukiwawczych. Przyjęty format opisu przekłada się bezpośrednio na możliwości mechanizmów wyszukiwawczych do zawężania instrukcji wyszukiwawczej. Innym sposobem pozyskiwania informacji w systemie Intute są narzędzia dostępu przedmiotowego. Umożliwiają one odnalezienie przez użytkownika odesłań hipertekstowych do poszukiwanych dokumentów za pomocą przyjętego do odwzorowania ich treści i/lub formy zbioru leksyki w układzie

hierarchicznym lub płaskim. Dostęp do informacji za pomocą tych narzędzi jest pochodną organizacji informacji, jaka została nałożona na zbiór wyszukiwawczy w systemie. Zbiór leksyki zawiera się najczęściej w słowniku wybranego języka informacyjno-wyszukiwawczego (JIW). Struktura hierarchiczna przejęta jest z organizacji pola semantycznego opracowanej dla danego JIW. Każde wyrażenie pochodzące z takiej struktury wyposażone jest w odesłanie hipertekstowe

- do wyrażen pozostających z nim w określonych relacjach (najczęściej są to relacje hierarchii tematycznej, generyczne lub mereologiczne),
- do charakterystyk wyszukiwawczych dokumentów, w których opisie je użyto.

Przeglądanie takiej struktury (ang. browsing) polega na ciągłym przeformułowaniu instrukcji wyszukiwawczej, aż do odnalezienia wyrażenia, które reprezentuje poszukiwane przez użytkownika zagadnienie czy też temat. W systemie Intute, każda z czterech grup tematycznych została zorganizowana za pomocą innych narzędzi. W tym celu posłużono się podzbiorami kilku słowników JIW oraz wybranymi fragmentami ich struktur paradygmatycznych. Np. w grupie tematycznej nauka i technologia, dział informatyka został zorganizowany za pomocą *ACM Computing Classification System*, w grupie tematycznej medycyna i nauka o zdrowiu, dział weterynaria został zorganizowany za pomocą *Klasyfikacji Biblioteki Kongresu* z wykorzystaniem słownictwa z *Tezaurusu CAB*. Narzędzia dostępu przedmiotowego są ważnym elementem procesu pozyskiwania informacji w dziedzinowych systemach hipertekstowych. Ich efektywność została potwierdzona przez wielu badaczy [m.in. Koch, Day, Neuroth 2004; Koch, Golub, Ado 2006], dlatego też są one postrzegane wraz z mechanizmami wyszukiwawczymi jako komplementarne narzędzia pozyskiwania informacji w tego rodzaju systemach.

Kompleksowość systemu Intute przejawia się również w jego dodatkowych funkcjach. Dotyczą one przede wszystkim szczegółowej specyfikacji usług informacyjnych, rozbudowanego systemu pomocy oraz szkoleń wirtualnych, a także personalizacji oferowanych usług. Są one istotne z punktu widzenia efektywności systemu oraz jego całościowej oceny przez użytkownika. Odpowiedni format opisu obiektów informacyjnych oraz edukacyjna perspektywa w odniesieniu do ich selekcji to dostosowanie systemu do potrzeb grupy użytkowników docelowych. Specyfikacja usług w kontekście systemu Intute polega na wprowadzeniu takich usług informacyjnych, które są potencjalnie pożądane przez studentów, nauczycieli i badaczy. Oprócz wspomnianych grup tematycznych Intute oferuje bogatą ofertę w tym zakresie. System udostępnia:

- bazę informacji o konferencjach i sympozjach naukowych w Wielkiej Brytanii,
- bazę czasopism elektronicznych,
- bazę teleadresową instytucji naukowych w Wielkiej Brytanii,
- możliwość uczestnictwa w kanałach komunikacyjnych RSS oraz blogach dotyczących poszczególnych grup tematycznych,

- system informacji o nowych rekordach oraz aktualnościach z życia nauki,
- kalendarium ważnych wydarzeń,
- przewodniki po 270 państwach uwzględniające informacje o ich historii, gospodarce oraz dane demograficzne,
- bazę danych o rynku pracy w Wielkiej Brytanii.

Korzystanie z wymienionych usług, tak jak z zasobów systemu jest nieodpłatne. Tym samym użytkownicy mają do dyspozycji ogólnie dostępne, kompleksowe narzędzie informacyjne dostosowane do ich potrzeb.

Personalizacja usług w Intute polega na umożliwieniu użytkownikowi utworzenia indywidualnego konta w systemie. W tym celu konieczne jest wypełnienie odpowiedniego formularza, który zawiera pytania dotyczące zainteresowań oraz określenia przynależności jednej z sześciu grup użytkowników (wykładowca, bibliotekarz, badacz, student, nauczyciel, inne). Przyporządkowanie do jednej z wymienionych grup jest pierwszym krokiem w celu zindywidualizowania dalszych usług i form interakcji z systemem. Posługiwanie się własnym profilem umożliwia rejestrację sesji wyszukiwawczych oraz otrzymywanych rezultatów wyszukiwania. System umożliwia zachowanie dowolnej instrukcji wyszukiwawczej wysłanej przez użytkownika wraz z rezultatami wyszukiwania, a także dowolnego rekordu odnalezioneego podczas innej formy pozyskiwania informacji. Zarządzanie tego typu danymi jest możliwe dzięki trzem funkcjom systemu:

- Moje rekordy (ang. my records)
- Moje wyszukiwania (ang. my searches)
- Moje tematy (ang. my subjects)

W każdym z paneli, gromadzone są informacje o różnej działalności użytkownika z wykorzystaniem narzędzi oferowanych przez system. Oprócz listy rekordów, instrukcji wyszukiwawczych czy interesujących użytkownika tematów, system udostępnia nieskomplikowane narzędzia do zarządzania tymi informacjami. W przypadku zbioru zgromadzonych rekordów istnieje możliwość opisania ich (tzw. znacznikowanie, ang. tagging, od tag - znacznik) za pomocą niekontrolowanych słów kluczowych oraz wyeksportowania ich charakterystyk w postaci kodu HTML, JAVA lub jako wiadomość e-mail. W panelu *Moje wyszukiwania* każde zarejestrowane zapytanie informacyjne użytkownika do mechanizmu wyszukiwawczego systemu jest uporządkowane chronologicznie wg daty jego wysłania. Instrukcje te są wyposażone w odesłania hipertekstowe, a ich aktywowanie powoduje ponowne wysłanie zapytania. Funkcja *Moje tematy* umożliwia zaznaczenie interesujących użytkownika tematów szczegółowych konstytuujących cztery główne grupy przedmiotowe. Spowoduje to gotowość systemu do selektywnej dystrybucji informacji o nowych rekordach opisujących dokumenty przyporządkowane do danych tematów dla odpowiednich użytkowników.

Oprócz pełnej personalizacji usług informacyjnych, system Intute udostępnia informacje przygotowane specjalnie dla wybranych grup użytkowników. W tym przypadku całość usług była tworzona pod kątem środowiska akademickiego (HE) Wielkiej Brytanii, a wydzieloną grupę użytkowników stanowi sektor kształcenia na poziomie poniżej szkolnictwa wyższego (FE). Ta grupa ma do dyspozycji podzbiór zbioru informacyjnego systemu, w postaci dokumentów elektronicznych adekwatnych dla potrzeb i poziomu intelektualnego jej członków.

Ciekawym rozwiązaniem zastosowanym w Intute jest rozbudowany system pomocy oraz szkoleń wirtualnych dla użytkowników systemu. Intute kontynuuje projekt zapoczątkowany przez RDN pod nazwą *Virtual Training Suite*. Jego celem jest kształtowanie kompetencji informacyjnych użytkowników, w tym przypadku internetowych usług informacyjnych opartych na modelu dziedzinowego udostępniania informacji. Szkolenia te mają formę samodzielnej pracy użytkownika z odpowiednim programem, który ma za zadanie przekazać mu wiedzę o systemie, wykształcić umiejętność formułowania problemu wyszukiwawczego i określić kryteria oceny jakości pozyskanych informacji. Efektywność tego projektu polega m.in. na dostosowaniu kursów do grup tematycznych systemu, którymi zainteresowani są użytkownicy. W ramach usług oferowanych przez Intute dostępny jest również program pt. *Internet Detective*. Program ten jest kontynuacją projektu, który funkcjonował pod koniec lat dziewięćdziesiątych ubiegłego stulecia [Worsfold, Hiom 1998]. Jego nowa wersja została oficjalnie uruchomiona wraz ze startem systemu Intute. Cel zarówno poprzedniej jak i aktualnej wersji jest ten sam, czyli wykształcenie w studentach umiejętności efektywnego wyszukiwania oraz racjonalnej oceny zasobów Internetu, które mogą być wykorzystane w procesach dydaktycznych [Place, Kendall 2006]. Projekt ten, podobnie jak *Virtual Training Suite* funkcjonuje w postaci przewodnika do samodzielnej pracy użytkownika z programem komputerowym. Dodatkową formą doskonalenia kompetencji informacyjnych użytkowników jest cykl seminariów organizowanych przez Intute. Od sierpnia 2006 roku odbyły się trzy tego rodzaju szkolenia:

- Internet w biznesie i turystyce (Internet for Business and Tourism), Bristol, 13 września 2006,
- Internet w nauce o zdrowiu (Internet for Health), Nottingham, 18 września 2006,
- Internet w naukach społecznych (Internet for Social Research), Birmingham, 20 września 2006.

Uczestnictwo we wszystkich seminariach było bezpłatne. Program szkoleń obejmował m.in. wiedzę na temat kryteriów oceny źródeł internetowych z poszczególnych dziedzin wiedzy, umiejętności ich wyszukiwania oraz selekcji.

Działalność popularyzatorska i dydaktyczna Intute to ważny element zarówno w komunikacji naukowej w Wielkiej Brytanii jak i kształtowaniu kompetencji informacyjnych w określonych grupach społecznych. Kolejną ważną płaszczyzną działalności Intute jest współpraca w ramach licznych projektów badawczych realizowanych przez jednostki naukowe w Wielkiej Brytanii.

Na tym polu, system ten odgrywa ważną rolę w adaptacjach nowych technologii szczególnie w projektach edukacyjnych. Do najważniejszych projektów zaliczają się:

- eBank UK – projekt finansowany przez JISC i prowadzony przez UKOLN, którego celem jest stworzenie repozytorium metadanych (z wykorzystaniem technologii Open Archives Initiative Protocol for Metadata Harvesting - OAI-PMH) z zakresu chemii, fizyki i zagadnień pokrewnych, które będzie zarządzane przez systemy edukacyjne (ang. learning management systems) oraz włączone w obieg informacji w nauce [Lyon 2003],
- INFORMS, projekt koordynowany przez w Uniwersytecie w Huddersfield, którego celem jest tworzenie specjalnego oprogramowania dla wirtualnych środowisk edukacyjnych. Celem tych narzędzi jest koordynacja procesów poznawczych towarzyszących zajęciom dydaktycznym z wykorzystaniem dokumentów elektronicznych na poziomie szkolnictwa wyższego.

Przedstawiona charakterystyka systemu Intute miała na celu określenie jego podstawowych funkcji oraz kierunków działalności w kontekście obiegu informacji w nauce w Wielkiej Brytanii. Jako jeden z największych tego typu systemów pełni on ważne funkcje szczególnie w nauce w Wielkiej Brytanii. Istnieje wiele systemów, które nie reprezentują tak rozbudowanej infrastruktury. Jednak działalność takich systemów jak Intute w Wielkiej Brytanii, eDna w Australii (www.edna.edu.au) czy SSG-FI (<http://www.sub.uni-goettingen.de/ssgfi/>) w Niemczech wyznacza kierunki rozwoju dziedzinowych systemów hipertekstowych. Rola tych systemów w kontekście zarządzania informacją w nauce polega m.in. na:

- transferze informacji do tego sektora,
- pełnieniu funkcji repozytoriów metadanych,
- dostosowaniu formatów opisu i narzędzi do potrzeb tej grupy użytkowników,
- zapewnianiu zaplecza programistycznego dla wirtualnych środowisk edukacyjnych.

Dziedzinowy model udostępniania informacji o wysokiej jakości zasobach Internetu odznacza się dużą efektywnością [HEERY, CARPENTER, DAY 2001], m.in. gdyż integruje mechanizmy wyszukiwawcze oraz narzędzia dostępu przedmiotowego. Tym samym znalazł on zastosowanie w elektronicznym obiegu informacji w nauce jako ważne narzędzie informacyjne.

Bibliografia.

1. DEMPSEY, LAW, D. *A policy context - eLib and the emergence of the subject gateways*. "Ariadne" 2000, nr 25.
2. HIOM, D. *Retrospective on the RDN*. "Ariadne" 2006, nr 47.
3. RUSBRIDGE, C. *After eLib*. "Ariadne" 2001, nr 26.
4. KNOWLTON, S. *A Preview of the New Cataloging Code*. "Michigan Library Association Forum" 2006 Wol.6, nr 1.
5. NAJJAR, J., DUVAL, E. *Actual Use of Learning Objects and Metadata*. "Bulletin of the IEEE Technical Committee on Digital Libraries" 2006 Wol.2, nr 2.
6. KOCH, T.; DAY, M.; NEUROTH *Searching and browsing multiple subject gateways in the Renardus service*. W: *Proceedings of the RC33 Sixth International Conference on Social Science Methodology*. Amsterdam, Stuttgart: 2004.
7. KOCH, T., GOLUB, K., ADO, A. *Users Browsing Behaviour in DDC-Based Web Services: A Log analysis*. "Cataloging & Classification Quarterly" 2006 Wol.42, nr 3/4.
8. WORSFOLD, HIOM, D. *Internet Detective*. "Ariadne" 1998, nr 18.
9. PLACE, E.; KENDALL, M. *Internet Detective - Back on the Case*. [Dokument elektroniczny] 2006, [dostęp: 2006-11-01]. Dostępny w World Wide Web: http://www.vts.intute.ac.uk/detective/downloads/press_release_Internet_Detective_June_2006.doc.
10. LYON, L. *eBank UK: Building the links between research data, scholarly communication and learning*. "Ariadne" 2003, nr 36.
11. HEERY, R., CARPENTER, L., DAY, M. *Renardus Project Developments and the Wider Digital Library Context*. "D-Lib Magazine" 2001 Wol.7, nr 4.
12. ROSZKOWSKI, M. *Źródła słownictwa kontrolowanego dla dziedzinowych systemów hipertekstowych w Internecie*. "Zagadnienia Informacji Naukowej" 2005, nr 2, s. 31-44.
13. DERFERT-WOLF, L. *Serwisy tematyczne o kontrolowanej jakości w Internecie - subject gateways*. "Biuletyn EBIB" 2004, nr 6.
14. CHAPMAN, A. *RDA: A New International Standard*. "Ariadne" 2006, nr 49.